DOCUMENT RESUME

ED 144 609 IR 005 257

AUTHOR Hjerppe, Roland
TITLE Experiences of an Interactive Retrieval System--ESRO/RECON.
INSTITUTION Royal Inst. of Tech., Stockholm (Sweden). Library.
REPORT NO TRITA-LIB-4035
PUB DATE Nov 74
NOTE 21p.; Tables 3 and 4 may be marginally legible due to type size

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Information Networks; *Information Retrieval; Information Science; *Libraries; *On Line Systems; Tables (Data)
IDENTIFIERS *Royal Institute of Technology (Sweden)

ABSTRACT

The library of the Royal Institute of Technology in Stockholm has been a node in the European network for on-line information storage and retrieval called ESRO/RECON since October 1972. Experiences gathered during these two years are described, especially on the problems encountered in conjunction with promotion of information services, problems of interaction with the on-line system, and problems of document back-up services. Proposals are given for development of features not available today but considered desirable for a fully interactive system. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED144609

REPORT TRITA-LIB-4035 NOVEMBER 1974

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

EXPERIENCES OF AN INTERACTIVE RETRIEVAL SYSTEM - ESRO/RECON

Roland Hjerppe

Royal Institute of Technology Library
S-100 44 Stockholm 70

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Stephan Schwarz

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

{Published in S. Schwarz (ed.): *The Interactive Library*, Stockholm 1975}

IR005257

EXPERIENCES OF AN INTERACTIVE RETRIEVAL SYSTEM - ESRO/RECON

ABSTRACT

The library of the Royal Institute of Technology in Stockholm has been a node in the European network for on-line information storage and retrieval called ESRO/RECON since October 1972.

Experiences gathered during these two years are described, especially on the problems encountered in conjunction with promotion of information services, problems of interaction with the on-line system and problems of document back-up services. Proposals are given for development of features not available today but considered desirable for a fully interactive system.

PREAMBLE

The ESRO/RECON is an on-line system for interactive search and retrieval of bibliographic information. Since all participants in the seminar will get a first hand experience of the system during the demonstrations and practical exercises this paper will not go into details of the system which is described comprehensively in (1, 2, 3, 4, 5).

The present ESRO/RECON is a modification and development of the NASA/RECON which in its turn traces its origins to the Lockheed DIALOG, a predecessor of which was known as CONVERSE. The system which is operated by the European Space Agency, ESA(formerly called European Space Research Organization, ESRO) through their Space documentation Service, SDS, has since March 1973 been implemented on an IBM 360/50 used as a dedicated computer for RECON. The computer and SDS are located in Frascati outside Rome. The network today encompasses 24 terminals in 8 European countries. The total number of references available for searching was by the end of October 1974 almost 3.5 millions, divided on 11 files. The largest of these files is CAC, Chemical Abstracts Condensates, with almost 1.5 million references, and the file with the longest time span is STAR+IAA, Scientific and Technical Aerospace Reports + International Aerospace Abstracts which start from 1962. A complete list of files is given in the paper by N. E. C. Isotta.

The Library of the Royal Institute of Technology has since October 1972 been connected to this network in order to investigate various aspects of interactive, retrospective information retrieval services. These investigations and the connection to ESRO/RECON are funded by the Swedish Council for Scientific Information and Documentation, SINFDOK.

At the time of writing this paper only the figures concerning the utilization and experiences during the first 13 months were available, the figures for the second year were being compiled. This paper will thus mainly report experiences relating to the first year, which in many senses was a transitory period and not wholly representative of a steady state.

PROMOTION

The promotion of information services can be divided into two related activities:

1 the use of various channels of communication for informing the potential users of the existence of the service.

2 the persuasion of the potential user that the service can A) fulfil his need, and B) do that better than any possible competitor.

The first task of informing the potential users of the availability of ESRO/RECON at the Library of the Royal Institute of Technology has been executed using a number of different communication channels.

- A leaflet containing general information about the system and its abilities was designed in co-operation with an advertising consultant who has designed all leaflets and other printed matters for the SDI-project. The new leaflet can thus from its appearance be recognized as emanating from the Library of the Royal Institute of Technology.

- This leaflet was distributed through the mailing lists of a number of professional societies and to all subscribers to the SDI-service. Approximately 3 000 leaflets were distributed during the first year.

- The announcement of the availability of the service was made in the form of a seminar in Stockholm at the library. Personal invitations to this seminar were sent through mailing lists of professional societies again, to all subscribers to the SDI-service and various selected persons. This seminar was attended by 260 persons. The SDI-service and the on-line system were at this occasion presented as two complementary services with the emphasis on RECON since the SDI-service has been the topic of many similar seminars. Another similar seminar was held on two consecutive days in Lund, and the terminal was temporarily transferred for that purpose. That seminar was called on by 250 persons. RECON was also shown at an one-day exhibition for 400 persons at the Technical Museum.

- Three daily newspapers carried articles on RECON in conjunction with the seminars.

- The RECON system and its abilities was incorporated as a feature in a daily news comment broadcast program by the Swedish Radio, called "Dagens Eko", which can be translated as "Daily Mirror".

- RECON has been announced and described in the newsletter of the Institute, "KTH-Nytt".

- The RECON system has been included in a leaflet, on publicly available information services in Sweden, which is produced by SINFDOK.

- RECON was demonstrated, after appointment, for individuals and small groups on 500 occasions. The total number of people attending these sessions is 1200. These demonstrations are of course to some extent a result of the other promotion activities.

Some indications of the efficiency of the various channels can be extracted from the questionnaire that each user receives together with the print-out of the results. One of the questions is: "How did you learn of the existence of RECON?" The number of users that have returned the questionnaire is quite small, 73, and the number of answers to this particular question is even smaller, n=65. Only very limited conclusions should thus be drawn from the figures in table 1 below.

TABLE 1. ANSWERS TO THE QUESTION: "HOW DID YOU LEARN OF THE EXISTENCE OF RECON?" IN FREQUENCY ORDER

Informed/recommended by another person	13
Informed by librarian/documentalist	12
From leaflet	12
By visits to/contacts with the KTH library	7
From internal information at KTH	6
As subscriber to the SDI-service	5
At seminar	3
From article/note	2
In library utilization course	2
Second time user	1
From SINFDOK-leaflet	1
Generally referred to the KTH library	1

The figures in table 1 seem to verify the two-step hypothesis for communication since more than half of those answering this question received their information from another person. The question "How did those persons informing the users become aware of the RECON system?" cannot however be answered without further investigations. One major conclusion that can be inferred from the figures is that all possible channels of communication should be used for disseminating information about the existence of the system, and that each channel will add a number of users which otherwise might have not been reached by the message.

The second task of persuading the user that RECON can fulfill his need, and do it better than the competitors has been somewhat easier since there is no comparable other service. The only other on-line retrospective service in Sweden is Medline which has an almost completely different user population. The only alternative to RECON is to do a manual search, and this has also been seen in the answers to another question in the questionnaire: "Why did you decide to utilize RECON?". Most of the few answers were " It´s the only available". This situation might however change if Lockheed or SDC were to market their systems in Sweden, since the files available through these systems are bigger and many more and since neither the communication costs nor utilization fees are prohibitive.

UTILIZATION

The files that were available, their starting year and the number of items on each file, at the beginning of connection, i. e. October 1972 and after one year in October 1973 are summarized in table 2.

TABLE 2. FILES AVAILABLE ON RECON AND NUMBER OF ITEMS ON EACH FILE

		Oct -72	Oct -73	
STAR+IAA	1962-	540 000	628 000	46%
Metadex	1969-	95 000	112 000	8%
Compendex	1969-	200 000	212 000	16%
GRA	1970-	85 000	100 000	7%
NSA	1968-	266 000	300 000	22%
Electronic Components	1971	8 000	12 000	1%
		1 194 000	1 364 000	100%

The utilization of the files and of RECON as a whole is displayed in table 3. A distinction has been made between a query and a search since one query, i.e. a problem or information need statement is usually matched against a number of files, a search thus being the search of one file for a query. Each query has on the average caused 3.3 searches to be made. In other words, the 211 commissioned queries have caused 690 searches. The number of queries can be taken as a measure of the utilization of RECON by the ultimate user, while the number of queries is a measure of the utilization of RECON by the searchers, documentalists. Almost all searches on RECON have been carried out by an intermediary, one of the documentalists at the SDI-service.

The routine is that the responsibility for the RECON system rotates on weekly basis among the documentalists. The one having responsibility for RECON during a specific week does not necessarily perform any of the searches that week since all of the documentalists have different professional background and since the one with the most pertinent education usually does the search. The documentalist responsible during a week rather ensures that RECON is utilized, and

disseminates the waiting queries among the documentalists. The query requestor is always invited to participate in the search and he also usually does so if he has the possibility. The documentalist will in that case, in a manner of speaking, be the pilot of the system while the requestor is the navigator. Our experience is that almost no customer has any desire to learn all the intricacies of the system since he is more interested in acquiring information than in the acquisition process. The information needed by someone doing research or development work is a means towards an end, the solving of a problem. The finding or retrieving of information is just a means for a means. Our experiences also quite clearly indicate that someone who does not regularly do searches on the system forgets many details that have to be looked up and which thus slows down the interaction.

The number of queries per month in table 2 shows a decline towards the end of the spring 1973, and then sharp rise in June. This sudden onset was due to the fact that RECON was offered free of charge to all researchers affiliated to KTH during June-August in order to get a higher utilization since only 9 queries were submitted in May, which meant that RECON was idle most of the time in May. This offer was also seen as a marketing effort which would raise the awareness of the system at KTH.

The figures in table 3 show that STAR+IAA and GRA have been the files most searched, each taking 27% of the searches and that Compendex is only a little less utilized with 24%. The file that has been utilized least is NSA with only 7% of the searches having been made on that file. When comparing these figures with the percentage of item per file in proportion to total number of items the difference with regard to NSA is even more striking. NSA has 22% of the items on

TABLE 3. NUMBER OF SEARCHES AND PRINTED REFERENCES PER DATABASE

	TOTAL						NASA			METADEX			COMPENDEX			GRA			NSA		
	Number of ref.	Number of queries	Average number ref/quer	Number of searches	Average number ref/search		Number of ref.	Number of searches	Average number ref/search	Number of ref.	Number of searches	Average number ref/search	Number of ref	Number of searches	Average number ref/search	Number of ref	Number of searches	Average number ref/search	Number of ref	Number of searches	Average number ref/search
Oct 72	1981	10	198	38	52		253	9	28	977	6	163	636	10	64	81	10	8	34	3	11
Nov	1409	8	176	25	56		436	6	73	706	7	101	202	7	29	65	4	16	0	1	0
Dec	3444	17	203	60	57		1702	16	106	517	8	65	458	16	29	363	16	23	404	4	101
Jan 73	3396	14	243	54	63		2136	14	153	179	8	22	395	12	33	644	14	46	42	6	7
Feb	1671	12	139	48	35		925	12	77	166	8	21	326	11	30	227	12	19	27	5	5
Mar	2877	10	288	33	87		805	8	101	221	4	55	1304	10	130	541	10	54	6	1	6
Apr	1194	6	199	17	70		228	5	46	333	2	167	418	6	70	215	4	54		0	
May	1200	7	171	27	44		335	7	48	257	4	64	225	6	38	379	6	63	4	4	1
Jun	6090	22	277	60	102		3010	19	158	412	3	137	961	17	57	1490	19	78	217	2	109
Jul	8971	29	309	88	102		4631	23	201	614	10	61	1543	19	81	1817	26	70	366	10	37
Aug	6301	30	210	87	72		3673	24	153	1637	15	109	357	20	18	630	24	26	4	4	1
Sept	7495	20	375	63	119		2849	16	178	3382	14	242	560	11	51	610	18	34	94	4	24
Oct	9050	26	348	90	101		3728	26	143	3230	16	202	1049	23	46	700	21	33	343	4	86
SUM	55079	211		690	(100%)		24711	185(27%)		12631	105 (15%)		8434	168 (24%)		7762	184 (27%)		1541	48 (7%)	
Ref. /Query																					
Average	261(100%)						117(45%)			60(23%)			40(15%)			37(14%)			7(3%)		
Average of averages	241																				
Ref. /Search																					
Average					80				134			120			50			42			32
Average of averages					74				113			108			52			40			32
Percentage of all files (from table 2)							46%			8%			16%			7%			22%		

RECON but has been utilized only for 7% of the searches and only 3% of the references printed out per query emanate from NSA. This is of course partly due to the fact that NSA is the most specialized of the files. GRA is the file with the highest utilization, 27% of the searches, in comparison with number of items, 7%, while the number of references printed per query is only 14%. Compendex has a similar utilization structure with 24% of the searches compared to 16% of the items and 15% of the printed references. This can be interpreted in the following way, GRA is a truly multidisciplinary file and very many of the queries will thus be matched with this file resulting in a large proportion of the searches, since the number of items is quite small and the time span short, not very many items will be found in every search. Since the number of searches is disproportionately high the number of references will still be high but not as high as the proportion of queries. The same interpretation can be applied, mutatis mutandis, to Compendex which is also multidisciplinary but not to the same extent as GRA.

The NASA file, STAR+IAA, on the other hand has almost half of the items available, covering a long time span. Since STAR+IAA is mission oriented file it might be called in some sense a "biased multidisciplinary" file, with the interpretation of the figures in table 3 that many of the queries, 27%, will because of the multidisciplinarity be searched against it, giving many references per query, 45%, since it is so large, 46% of the items.

The figures for Metadex might be interpreted as follows: Metadex is a "heavily biased multidisciplinary" file, with the result that a small number of the searches, 15%, will be made on that file. These searches will on the other hand be for queries especially suitable for that file and will thus find a high proportion of the references per query, 23%, considering the size of the file, 8%.

TABLE 4. DISTRIBUTION OF QUERIES (FREE/PAID, INDUSTRY/UNIV./OTHERS) AND TIME DELAYS

	NUMBER OF COMMISSIONED QUERIES						DELAY IN DAYS SUBMISSION-SEARCH			DELAY IN DAYS SEARCH-DISPATCH			DELAY I DAYS SUBMISSION-DISPATCH			SEARCHING TIME (HRS)		
	SUM	FREE	PAID	INDUSTRY	UNIV	OTHERS	SUM	NUMBER OF QUERIES	AVERAGE NUMBER DAYS/QUERY	SUM	NUMBER OF QUERIES	AVERAGE NUMBER DAYS/QUERY	SUM	NUMBER OF QUERIES	AVERAGE NUMBER DAYS/QUERY	SUM	NUMBER OF QUERIES	AVERAGE NUMBER HRS/QUERY
Oct 72	10	1	9	7	1	2	89	9	9.9	62	10	6.2	132	9	14.7	43.4	9	4.8
Nov	8	0	8	3	4	1	43	5	8.6	38	8	4.8	67	5	13.4	19.4	8	2.4
Dec	17	6	11	9	3	5	151	12	12.6	216	16	13.5	248	11	22.5	48.5	16	3.0
Jan 73	14	0	14	7	3	4	99	12	8.3	75	14	5.4	162	12	13.5	46.6	14	3.3
Feb	12	0	12	4	4	4	58	9	6.4	123	12	10.3	159	9	17.7	42.0	11	3.8
Mar	10	0	10	4	4	2	72	10	7.2	84	10	8.4	156	10	15.6	34.0	10	3.4
Apr	6	0	6	3	3	0	32	6	5.3	49	6	8.2	81	6	13.5	8.6	4	2.2
May	7	0	7	2	3	2	61	7	8.7	66	7	9.4	127	7	18.1	13.5	6	2.3
Jun	22	18	4	0	13	4	175	21	8.3	120	22	5.5	288	21	13.7	53.5	22	2.4
Jul	29	21	8	2	21	6	45	11	4.1	127	29	4.4	87	11	7.9	58.0	24	2.4
Aug	30	19	11	10	16	4	160	11	14.5	198	30	6.6	239	11	21.7	58.8	25	2.4
Sep	20	9	11	6	11	3	146	9	16.2	149	18	8.3	210	9	23.3	39.3	17	2.3
Oct	26	7	19	10	10	6	237	14	15.9	190	26	7.3	352	14	25.1	58.1	22	2.6
SUM	211	81	130	67	101	43	1368	136		1497	208		2308	135		523.7	188	
Percentage	100	38	62	32	48	20												
Average							10.1			7.2			17.1			2.8		
Average of averages									9.8			7.6			17.0			2.9

These interpretations are of course open to much discussion, but they seem to be supported by the figures given for average number of references per search. The average number of references per search for all searches is 80, with the average of averages being 74 which probably is closer to the median, while the corresponding figures for STAR+IAA and Metadex are 113 and 108 respectively. The figures for Compendex, GRA and NSA are 52 , 40 and 32 respectively. These figures, when compared with the average number of references per query show a marked discrepancy for Metadex, 108 against 60, which seems to support the interpretation put forward for Metadex, that many of the searches made on Metadex have been well suited for that file. The discrepancy for NSA, 32 against 7, is seen to be due to the fact that there have been a few queries well suited for NSA, searches made in December 1972 and June and September-October 1973. Most of the rest of the searches on NSA have yielded very little, quite probably because they were made more as a sort of check than with the hope of finding much.

Table 4 shows the distribution of queries with divisions for free, paid, industry, university, others. The effect of the free of charge offer to KTH affiliated can be seen quite clearly in this table. We can also see that a large proportion, 62% of the queries were paid for during this first year. The demonstrations mentioned earlier in conjunction with promotion are not included in tables 3 and 4. We can also see that queries originating from universities or colleges make up almost half of the queries and that others, mainly agencies and research institutions, do not differ much from industries as users of RECON.

In table 4 are also shown the delays, from submission to search and from search to dispatch and also the total delay, submission to dispatch. To the total delay should be added

one day for postal service within Sweden. This would mean that the total delay, i. e. the amount of time that has passed since the query was submitted to that moment when the result is received, is on the average 18 days. The average search time is in comparison almost 3 hours.

These delays can of course be explained but are nevertheless real to the customer. The delay between search and dispatch is due to the fact that all results exceeding approximately 10 references are printed in Frascati outside Rome. The postal service in Italy is nowadays notoriously ineffecient, for a variety of reasons. In order to assure that the results are not lost SDS has had to take the results twice weekly by courier to Paris where they have been mailed.

The delays between submission and search have two main causes. The first cause is that since many customers want to participate in the search, an appointment has to be made that suits both the customer and the system. The result is of course that the search is postponed to a date that suits both. This explanation is naturally not valid for those queries in which the customer cannot or does not want to participate in. Unfortunately there are no figures available for the proportions between these two situations.
The second cause resides in the fact that the search time per query is almost 3 hours. This means that more than two queries cannot usually be executed every day, especially when down-time is considered. Since there was no "search save" available during this period a remaining hour or hour and a half by the end of the day could not always be used for queries prognosticated to be time consuming. These factors together had the result that unprocessed queries piled up quite rapidly, especially if the system had a serious break-down. The situation this fall, 1974, has been that we have for at least two months had a queue of 10-20 queries waiting, and

this backlog is very hard to eliminate as long as the queries keep coming at a regular pace, which is what we in other circumstances would want them to do. The situation has been somewhat relieved by the terminal in Lund, but not enough.

The question naturally arises, "Can the total search time be shortened?". This does of course depend of the goals of and with the service. During the first two years great emphasis was put on coverage with the result that the searches performed have been comprehensive and thorough, as can be seen from the total searching time. One factor prolonging the search time was the fact that neither INSPEC nor CAC have been available although many of the queries could have been satisfied by only one of these two. The effect has been that the other files available have been overutilized in efforts at trying to give the customer something. The evaluation of the number of references received by the customer has shown that the documentalists, especially when searching alone, have a disposition towards making the searches to extensive, with a low degree of precision, 17% of immediate interest and 20% of secondary interest, n=29.

EVALUATION BY CUSTOMERS

As has been stated earlier, each customer receives with the results a questionnaire which he is asked to answer and return to KTH. The response rate has been very low so the figures in table 5 should only be taken as indications. As always we do not know what the great majority, those who have not answered, think of the system.
The figures indicate that almost every third customer thinks that he received too many references and that half of the customers were satisfied. A very high proportion of the customers, 70%, consider that the original documents lived

up to their expectations and that they had had definite use of the information. A very high proportion, 86% and 94% of those answering the questionnaire were going to use RECON again, respectively recommend it to colleagues. This is in many senses one of the ultimate tests. Oddly enough more of the customers considered recommending it than using it again themselves.

TABLE 5. ANSWERS TO SOME OF THE QUESTIONS IN THE QUESTIONNAIRE

		Number of answers		Total number of answers
Have your expectations of RECON been fulfilled	Yes	37	60%	61
	Doubtful	10	16%	
	No	14	24%	
Evaluation of number of references received	Satisf.	29	49%	59
	Too many	17	29%	
	Too few	13	22%	
Were the original documents as you expected	Yes	30	70%	43
	Doubtful	6	14%	
	No	7	16%	
Have you had any definite use of the information	Yes	31	70%	44
	Doubtful	3	7%	
	No	10	23%	
Are you going to use RECON again	Yes	45	86%	52
	No	7	14%	
Will you recommend RECON to your colleagues	Yes	50	94%	53
	No	3	6%	

ACCESS TO ORIGINAL DOCUMENTS

Since the SDI-service at the Library of the Royal Institute of Technology has been operational for many years, serving a large number of customers, the procedures for ordering copies of the original articles are running well. The customers of RECON have not experienced many problems, as long as their orders have been for journal articles. The problems arise in conjunction with the reports from STAR and GRA since they are seldom available in Sweden. This means that they often have to be bought from NTIS and this is a problem because the library that can provide copies of articles cannot function as a report broker for all the customers that are unaccustomed to the procedures for ordering reports from the various depositories. The resources of the library are too scarce to facilitate but the most scant assistance in these matters. Another impeding factor is the fact that these individual users´ orders will not receive the same attention at the depositories as a regular customer, and that they in some cases first have to write and ask for special order forms or coupons.

CONCLUSION

The discussions of features desired in truly interactive retrieval system that was promised in the abstract cannot be accomodated within the limits put on this paper and must thus remain and become the topic of a separate future paper. As a summing up it might be noted that the features noted by the customer are not its speed but rather the possibility for modifying and reiterating the search and the convenience in comparison with manual searches. The negative features are the short time spans covered and the need for repeating searches on different files.

REFERENCES:

(1) Isotta NEC
Aerospace information service-progress with the ESRO/ELDO computerised information network.
New Developments in Storage, Retrieval and Dissemination of Aerospace Information.
Neuilly sur Seine, Nato, AGARD, December 1973
AGARD Conference Proceedings no. 136

(2) Romerio G
The ESRO real time information network.
Technology Ireland, 4 (12) March 1973, pp 18-20, 25

(3) Raitt DI
Space Documentation Service: Operations Handbook
European Space Research Organisation, May 1971
ESRO SP-63.

(4) Walker DE, Ed.
Interactive Bibliographic Search: The User/Computer Interface.
AFIPS Press, Montvale NJ, 1971.

(5) Lancaster FW, Fayen EG
Information Retrieval On-Line
Melville Publishing Company
Los Angeles Calif., 1973